Can intergroup contact affect ingroup dynamics? Insights from a field study with Jewish and Arab-Palestinian youth in Israel

Ditlmann, Ruth K.; Samii, Cyrus

Postprint / Postprint
Zeitschriftenartikel / journal article

Ditlmann, Ruth K.; Samii, Cyrus

Article — Accepted Manuscript (Postprint)

Can intergroup contact affect ingroup dynamics? Insights from a field study with Jewish and Arab-Palestinian youth in Israel

Peace and Conflict: Journal of Peace Psychology

Provided in Cooperation with:
WZB Berlin Social Science Center

Suggested Citation: Ditlmann, Ruth K.; Samii, Cyrus (2016) : Can intergroup contact affect ingroup dynamics? Insights from a field study with Jewish and Arab-Palestinian youth in Israel, Peace and Conflict: Journal of Peace Psychology, ISSN 1532-7949, American Psychological Association, Washington, DC, Vol. 22, Iss. 4, pp. 380–392, http://dx.doi.org/10.1037/pac0000217

This Version is available at:
http://hdl.handle.net/10419/209584

Peace and Conflict: Journal of Peace Psychology

Manuscript version of

Can Intergroup Contact Affect Ingroup Dynamics? Insights From a Field Study With Jewish and Arab-Palestinian Youth in Israel

Ruth K. Ditlmann, Cyrus Samii

Funded by:

- United States Agency for International Development, United States Institute for Peace

Abstract

How can intergroup contact programs affect conflict-ridden communities besides improving the out-group attitudes of participating individuals? We address this question by examining the effects of an intergroup contact intervention on *in-group* dynamics that may mitigate inter-group conflict. We also examine how outgroup attitudes and psychological resources mediate such effects. We present the results from a difference-in-differences design with 149 Jewish and Arab-Palestinian youth, some of whom participated in an intergroup contact and sports peace program operated by an NGO in Israel. Our main outcome is one's tendency to *censure* ingroup members' provocations toward the outgroup. As expected, we find a positive impact of the program on ingroup censuring. However, this result is only marginally significant. We find a positive effect of program participation on outgroup attitudes among Jewish youth as expected. To our surprise, among Arab-Palestinian youth, we find a negative effect on outgroup attitudes. Exploring the underlying processes and group-based differences further, we find that outgroup regard mediates the effect of intergroup contact on ingroup censuring for Jewish youth. We find no evidence for mediation among Arab-Palestinian youth but a positive association between ingroup censuring and psychological resources. These results suggest that the psychological conditions of ingroup censuring may differ by group. We discuss implications for peace-building interventions in societies with groups in conflict.

Keywords: Ingroup censuring, intergroup contact, outgroup regard, psychological resources

Can Intergroup Contact Affect Ingroup Dynamics? Insights from Jewish and Arab-Palestinian Youth in Israel

Every year governments, non-governmental organizations and companies spend millions of dollars on "people-to-people" peace-building programs or intergroup contact interventions in conflict-ridden societies (e.g., Peacebuilding and reconciliation, 2014; Football for hope, 2014). The assumption behind many of these programs is that they will help transform conflict-ridden into peaceful societies. Yet, little is known about the effectiveness of these programs in societies at conflict and if they go beyond improving the outgroup attitudes held by immediate program beneficiaries. The current research examines if and how an intergroup contact intervention in Israel affects the way participants relate to *ingroup* peers in ways that mitigate conflict with the outgroup. We further explore how intra-group and inter-group dynamics relate.

This paper draws on research with an NGO that operates a peace program with Jewish and Arab-Palestinian residents of Israel. By collaborating with an NGO we maximize the external validity of our research. First, we use a "treatment" that actually serves youth across Israel. Second, our research includes underprivileged Jewish and Arab-Palestinian youth in Israel. These groups are strongly affected by the socio-political context, yet rarely participate in psychological research, probably because researchers lack access to their communities. Working with such populations is an important step for expanding the reach of social psychology (Heinrich, Heine, & Norenzayan, 2010).

Inter-group Contact in Societies at Conflict

A meta-analysis of 515 studies with over 600 samples found that intergroup contact is associated with more positive attitudes toward outgroups, especially among members of majority groups in society (Tropp & Pettigrew, 2005). Yet, out of those 515 studies less than 3% percent involved groups in conflict. Indeed, in societies of ongoing conflict, positive attitude change is especially difficult to achieve (Hameiri, Bar-Tal, & Halperin, 2014). Arguably, these same societies would have the most to gain from successful intergroup contact interventions.

In Israel the evidence regarding intergroup contact programs is mixed. Recent pre-post evaluations of a co-ethnic soccer program, other co-ethnic sports programs, and the peace education program "Seeds for Peace" showed significant improvements in the outgroup attitudes of Jewish-Israelis, Arab-Palestinian citizens of Israel, and Palestinians (Galily, Leitner, & Shimon, 2013a; Galily, Leitner, & Shimon, 2013b; Leitner, Galily, & Shimon, 2012; Schroeder & Risen, 2016). Other research is less optimistic about intergroup encounters as a tool for promoting conflict resolution in Israel. A qualitative study revealed that after participating in a peace program, the outgroup attitudes of Jewish Israelis, Arab-Palestinian citizens of Israel, and Palestinians were even more polarized than at baseline (Hammack, 2006). In another study, in two out of four groups in an inter-ethnic program that provided peace education and training for soccer coaches, the Palestinian-Arabic coaches left early (Livtak-Hirsch, Galily, & Leitner, 2016). Finally, Amir et al. predicted that outgroup attitude changes for West Bank Palestinians should be negative because West Bank authorities opposed intergroup contact, and found

a (non-significant) trend in the predicted direction (Amir, Bizman, Ben-Ari, & Rivner, 1980).

Critical scholarship proposes that if they reflect the power asymmetries between the Jewish majority and the Arab-Palestinian minority in society, intergroup encounters in Israel are not effective for promoting conflict resolution (Maoz, 2000a; Maoz, 2000b). Furthermore, for Arab-Palestinians in Israel, as for minority groups in other societies, intergroup contact can reduce the motivation to work for social change (Dixon, Levine, Reicher, & Durrheim, 2012; Saguy, Tausch, Dovidio, & Pratto, 2009). In our view, the mixed results necessitate further empirical research that considers program effects in general and also separately by ethnicity, as well as outcomes beyond outgroup attitudes.

Ingroup censuring

Traditionally, much research has examined improved outgroup attitudes as the primary outcome of intergroup contact interventions. Yet, even if contact effectively reduces prejudice among groups in conflict, only a small number of people participate in intergroup contact programs. For the program to have effects beyond the immediate participants these participants need to influence members of their community.

One way to influence one's ingroup peers is by *censuring* their aggressive actions toward outgroup members. Individuals who censure other ingroup members use intra-group conflict and dissent to promote inter-group cooperation, a strategy that has received little attention in social psychology despite its importance for transforming societies (Cikara & Paluck, 2013). Censuring ingroup members' aggression toward the outgroup is a form of "ingroup policing," a concept from the political science literature

on inter-group cooperation (Brubaker & Laitin, 1998; Fearon & Laitin, 1996). Indeed, Fearon and Laitin (1996) propose that historically, policing of the ingroup has been an essential mechanism through which ethnic groups have maintained cooperative relations with each other. Ingroup censuring also provides a clear way through which individuals can affect group dynamics, and therefore through which individual-level interventions might transform communities. This raises the question of whether peace-building programs can affect ingroup censuring inclinations.

In everyday conflicts, ingroup censuring takes the form of publicly or privately admonishing the hateful comments or aggressive actions of an ingroup member towards the outgroup. This behavior is illustrated nicely in the following quote from a participant in the co-ethnic sports league that forms the basis of the peace program studied below:

> We had a big game in the [co-ethnic] league, and one of the parents from the other team got mad and started making racial comments, and my mom was at that game, and she got so mad at him. She started telling him: You don't talk to my girls like that. They are all my girls, and you won't say anything like that to them.

The peace program we study focuses on ordinary youth in Israel. As such, we examine censuring in everyday intergroup conflicts that might occur when, for example, people attend a sports game or argue over a seat on a public bus. Nevertheless, many examples are available when ingroup censuring affected the Israeli-Palestinian conflict on a larger scale. For example, in the highly publicized August 2014 Arab-Jewish wedding of Mahmud Mansur and Morel Malka near Tel Aviv, Jewish Israelis stood up to rightwing extremist Jewish Israelis over the right of Jews and Arabs in Israel to marry

(Rousseau, 2014, August 18). Similarly, when an Arab resident provocatively blasted loud music from his car during Yom Kippur observances, Arab leaders from Acre resolved resulting tensions between Arab-Palestinians and Jews by condemning actions of the provocateur (Al Jazeera, 2008, October 13). Even though our research looks at ingroup censuring in everyday conflict, these examples illustrate that this is an outcome of high relevance for the conflict at large.

The primary objective of the current research is to study the effects of a peace program in Israel on ingroup censuring, an intra-group outcome. Given the mixed evidence regarding the effectiveness of intergroup-contact programs in Israel so far (Amir et al., 1980; Galily et al., 2013a, 2013b; Hammack, 2006) a secondary objective is to examine the programs' effects on outgroup attitudes, a traditional, inter-group outcome. A third and final objective is to bring intra- and inter-group outcomes together by exploring the psychological processes that underlie ingroup censuring.

We consider two possible motivating factors: (1) outgroup regard and (2) psychological resources. The first is based on the idea that higher inclinations to censure ingroup members who take aggressive action toward the outgroup may reflect higher regard for the outgroup. Higher outgroup regard is associated with greater empathy towards outgroup members who are experiencing harm (Avenanti, Sirigu, & Aglioti, 2010), and with valuing inter-group harmony more (Allport, 1979). Presumably such associations are the reason that research on inter-group contact and peace-building programs tends to focus on inter-group outcomes, such as measures of social distance or affect toward outgroup members.

The second factor we consider is based on the idea that higher inclinations to censure ingroup members reflect greater access to psychological resources such as self-efficacy or self-esteem (Bandura, 1982). Dissenting from one's ingroup is often non-normative behavior (Cikara & Paluck, 2013), meaning that it could be costly if it triggers disapproval from other ingroup members. Access to psychological resources such as self-efficacy and self-esteem may be necessary for someone to be willing to censure ingroup members when faced with such costs. Our analysis below examines the effects of the program on these two motivating factors, the structural relationship between these motivating factors and ingroup censuring, and an analysis of the extent to which these factors mediate any effect of the program on ingroup censuring.

Research Context: An Interethnic Sports Program with Jewish and Arab-Palestinian Youth

We collaborated with an NGO that offers a program involving inter-group contact in a co-ethnic sports league and a peace curriculum to Jewish and Arab-Palestinian youth in different locations throughout Israel. The youth practice a team sport weekly in segregated teams, and meet approximately once per month with a partner outgroup team for joint practices. In addition, youth learn non-violent forms of conflict resolution and humanization of the other through ball drills, role-playing, and games. Each program year starts in the fall and ends in the spring.

As much as possible, the NGO establishes the conditions of "optimal" intergroup contact (Allport, 1979). The youth have a shared goal of succeeding at their sport, and team sport requires high levels of cooperation. During shared practices coaches form

mixed-ethnic teams and run drills that are only successful if all participants cooperate, e.g. passing a ball between two people. Authority support is available in the form of support from parents and coaches. To ensure status symmetry the NGO makes sure that youth come from similar socio-economic backgrounds, have a similar level of athletic skills, and that a similar share of coaches and management are Jewish and Arab-Palestinian. Our data on Jewish and Arab-Palestinian program locations suggests similar rates of unemployment, which reflects the NGO's effort to align class positions (see Appendix A). Because this peace program fulfills all of the optimal conditions for intergroup contact, we expected a positive impact of program participation on ingroup censuring and on outgroup attitudes. Moreover, because the program attempts to build character and skills, we expected a positive effect on psychological resources.

Our understanding of the context also led us to expect ethnicity to play a moderating role. As described above, the intergroup contact literature suggests that intergroup contact affects members of majority / high status groups and members of minority / low status groups differently (Tropp & Pettigrew, 2005). Recent social psychology literature on asymmetric power relations (e.g., Maoz, 2012; Saguy et al., 2009; Saguy, Chernyak-Hai, Andrighetto, & Bryson, 2013) conceptualizes Jewish Israelis as a high status group and Arab-Palestinian residents of Israel a low status group. Even though Jewish and Arab-Palestinian participants in our study come from communities with similar rates of unemployment (see appendix A.1), the conceptualization as high and low status groups is consistent with descriptive evidence

regarding the economic situation of the groups in general terms.[1] For example, the income of Arab-Palestinian wage earners was 63.5% the amount earned by Jewish wage earners in 2013, and 67.2% in 2014 (Bassok & Heruti-Sover, 2015), and Arab-Palestinians were rejected at a higher rate (30.5%) than Jews (17.3%) when applying for a first university degree in Israel (Central Bureau of Statistics State of Israel, 2014). We thus include participants' ethnicity as a moderator to see if the program effects differ for Jewish and Arab-Palestinian subjects.

Empirical Methods

Research Subjects

We collected data from 72 Palestinian and 77 Jewish subjects[2] in Israeli-Jewish and Arab-Palestinian communities in Israel between January and June 2013. The data was collected as part of an ongoing multi-year, multi-site study on the effects of the peace program described above. Our sample includes members of the new incoming 2013 cohort of participants in the program (N = 53), a sample of non-participant peers recruited from the same schools and community centers as the participants (N = 59), as well as a group of veteran participants (N = 37). The sample of incoming cohort members and veteran participants was the product of the collaborating NGO's normal recruitment practices, which include flyers, word-of-mouth, and special events conducted in local

[1] The assignment of groups to high versus low status could vary depending on whether one defines status in terms of allies in the broader region, perceived threats to personal safety, or historic precedence of victimization. Indeed, groups' roles as perpetrators or victims, often (although not necessarily) connected to their status, are often contested in conflicts (Noor, Shnabel, Halabi, & Nadler, 2012). In this article, we define status in terms of access to material goods and resources.

[2] We use the term "subjects" to refer to youth who participated in the research and to distinguish from those who participated in the contact and sports peace program, that is "participants".

schools and community centers. The non-participant peer sample was constructed by approaching the same schools and community centers and having teachers and community center staff identify youth with the same background as participant youth, but who had not participated in the program for incidental reasons.

Table 1 shows how the data for Jewish and Arab-Palestinian subjects are distributed over key background characteristics. The sample is predominately female. In our analyses reported below, we ensure that our results are robust to balancing the gender composition across the participant and non-participant peer samples by excluding males from the analysis. A minority of Jewish subjects comes from immigrant households; a similar share of Arab-Palestinian subjects comes from households with parents that had resettled from elsewhere in Israel, Palestine, or other Arab countries (see D.2 in Appendix D for details on families' places of origin). Subjects' (including veterans) ages range from 10 to 22, with the majority being between the ages of 11 and 14. The incoming program cohort and non-participant peer group cohort ages range from 10 to 15, although the participant youth are slightly younger on average (Table D.1 in Appendix D). We also control for age in the analyses below. Arab-Palestinian subjects report more religiosity on average than their Jewish counterparts (the scale is scored such that 1 means most religious).

[Table 1]

All parents were informed about our study and gave consent for their children. The parents of participants learned about the study and gave consent as part of a program registration form they sign before the beginning of each season. The NGO sent information about the research to parents of subjects in the control group. All youth were

allowed to participate in the research only if their parents gave consent. If they had questions about the research parents could call our local research coordinator, who spoke Hebrew and Arabic.

Measurement Instruments

The outcome variables of interest are ingroup censuring, outgroup regard, and psychological resources. To measure ingroup censuring we created a set of scenarios that describe how a member of the ingroup (e.g., a Jewish youth if subject is Jewish) aggresses against a member of the outgroup (e.g. an Arab-Palestinian youth if the subject is Jewish). For example, in one scenario an ingroup youth aggressively pushes an outgroup youth out of the way for easier access to the candy shelf in a supermarket. (See Appendix B for all scenarios). After reading each scenario, subjects had an opportunity to censure the action of the ingroup aggressor using two 4-point Likert scales, each asking some variation of "How okay was it for X to shove/scream at/use derogatory words with Y?", with answer choices ranging from 1 (*it's perfectly okay*) to 4 (*it's really wrong*). Two censuring opportunities were used for each scenario to increase measurement validity. Higher scores are taken to mean higher inclination to censure ingroup members' aggression toward the outgroup. As can be seen in Table 1, average ingroup censuring index scores are nearly identical for Jewish and Arab-Palestinian subjects.

In relating this measure to our theoretical motivation, one may wonder if such private censuring is associated with tendencies to censure publicly. We show in Appendix B that our private censuring measure has a strong positive correlation with self-reported frequency of having defended outgroup members' points of view vis-à-vis

friends and family. This is indicative of the measure's validity in capturing willingness to take a stand against ingroup members on the outgroup's behalf. That being the case, the cost to public ingroup censuring is high in the current political climate in Israel (Zonszein, 2014, September 26). While observing public censuring would be ideal for the purposes of our analysis, encouraging our subjects to publicly ingroup censure would pose a threat to their safety, and we therefore restricted measurement to private censuring. To maximize proximity to public censuring behavior, we modeled our ingroup censuring measure after "beliefs about aggression" scales that correlate positively and substantially with behaviors in the classroom (Möller & Krahé, 2009; Huesmann & Guerra, 1997).

In January 2013, our first wave of data collection, Jewish subjects evaluated four scenarios of ingroup aggression against an outgroup member, and Arab-Palestinian subjects evaluated three such scenarios. Jewish and Arab-Palestinian subjects' responses to all scenarios correlated highly ($\alpha = 0.92$, $\alpha = 0.86$ respectively). In light of this, in the second wave of data collection in June 2013 we shortened the survey by presenting only one or two of these scenarios to each subject (depending on random assignment). For subjects receiving two scenarios in the second wave, Cronbach's alpha coefficient (α) is 0.80, and for subjects participating in both waves of data collection, the over-time correlation is 0.61 ($p < .001$). This indicates that shortening the instrument did not undermine validity. In our analyses, we use the simple averages across all scenarios of ingroup on outgroup aggression presented to subjects in each wave. In the regression analyses below, the outcome is a standardized version of this variable, with the

standardization done against the pooled mean and standard deviation. The regression coefficients are effect sizes at the scale of outcome standard deviations.

We measured outgroup regard with two instruments: a feeling thermometer and a social distance scale. The feeling thermometer is a widely used measure of affect towards outgroups (e.g., Paolini, Hewstone, Cairns, & Voci, 2004). Arab-Palestinian subjects indicated their feelings toward "Jews" and "Israelis" on a "feeling thermometer" anchored by 0 (*very cold)*, 50 (*neither cold nor warm*), and 100 (*very warm*). Jewish subjects indicated their feelings towards "Arabs" and "Muslims" using the same scale. The social distance scale is a widely used measure of how close of a relationship subjects can imagine having with someone from the other group (Bogardus, 1933; Parrillo & Donoghue, 2005). We modified some of the items for age appropriateness. Subjects indicated their agreement with each of 6 items that vary in closeness of relationship (e.g., "I would be willing to study in the same school as members of the other group", "I would be willing to invite members of the other group to my house") on a 5 point Likert scale from 1 (*strongly disagree)* to 5 (*strongly agree*).

We aggregate the information from the feeling thermometer and the social distance items by conducting a factor analysis. We construct indices by using a principal factors model with regression scoring on the unrotated solution. Even though we used distinct scales from psychological research the results of the factor analysis indicate one dominant factor for the outgroup regard variables. (A scree plot is shown in Appendix C.) We use the score from the dominant factor in our analyses to capture variation in outgroup regard. As can be seen in Table 1, Arab-Palestinian subjects report slightly

higher outgroup regard than Jewish subjects, but the difference is not statistically significant at conventional levels.

We measured psychological resources using instruments for state self-esteem, collective self-esteem, and self-efficacy. These three constructs constitute a general category of psychological resources because they have been shown to contribute to higher wellbeing (Crocker, Luhtanen, Blaine, & Broadnax, 1994) and to the propensity to take action to improve a situation (Bandura, 1982). State self-esteem measures how positively individuals evaluate themselves in a given context (Heatherton & Polivy, 1991). All our subjects indicated their level of agreement with each of 20 items (e.g., "I feel displeased with myself") on a Likert scale from 1 (*strongly disagree)* to 7 (*strongly agree*). The collective self-esteem scale assesses perceived public regard of one's group (Luhtanen & Crocker, 1992). Our subjects indicated their level of agreement with each of eight items (e.g., "In general, others respect my group") on a Likert scale from 1 (*strongly disagree)* to 7 (*strongly agree*). Finally, self-efficacy measures individuals' faith in their ability to implement action plans (Sherer & Adams, 1983). Our subjects, again, indicated their level of agreement with each of 17 items (e.g., "When I make plans, I am certain I can make them work") on a Likert scale from 1 (*not at all true)* to 7 (*completely true*).

As with the outgroup regard index, we aggregate the information from the psychological resource variables by conducting a factor analysis and construct an index by using a principal factors model with regression scoring on the unrotated solution. Even though we used distinct scales from psychological research the results of the factor analysis indicate one dominant factor for Jewish and Arab-Palestinian youth for the state

self-esteem, collective self-esteem and self-efficacy items. (Scree plot is shown in Appendix C.) We use the score from the dominant factor in our analysis to capture variation in psychological resources. As can be seen in Table 1, Arab-Palestinian subjects report slightly higher psychological resources, but the difference is not statistically significant at conventional levels.

Procedures and Analytical Methods

We used a "difference-in-differences" design to estimate the causal effect of the peace program between January and June 2013. This involves comparing the change in outcomes over time among participants (the first "difference") to the change in outcomes over time among non-participants (the second "difference"), with the program effect estimated as the difference between these two changes (the "difference-in-differences"). The difference-in-differences design accounts for baseline differences between participant and control groups that would otherwise bias a simple cross-sectional comparison; it also explicitly accounts for over-time trends that would otherwise bias a simple before-after comparison (Angrist & Pischke 2009, Ch. 5). So long as the outcome trends for program and control group participants would be similar under circumstances where no program was available (the "parallel trends under control" assumption), a difference-in-differences design estimates a valid causal effect. Estimating a causal effect is crucial when investigating intergroup contact interventions to avoid the potential participant selection bias that may confound positive correlations between contact and outcomes of interest (Pettigrew, 1998).

We carry out the difference-in-differences analysis using only data from the newly incoming cohort members and their non-participant peers. This sample includes

61 Jewish subjects, of which 37 completed both waves of research, 17 completed only wave 1, and seven completed only wave 2, and 51 Arab-Palestinian subjects, of which 26 completed both waves, 21 completed only wave 1, and four completed only wave 2. Regardless of whether they completed one or two waves of research, all subjects in the intervention participated in all the activities described above. To preserve their anonymity, we assigned age-appropriate and easy-to-remember passwords to youth (e.g., famous athletes) to link their wave 1 and wave 2 responses.

To compute the difference-in-differences estimates, one uses ordinary least squares (OLS) regression, regressing the outcome on indicators for whether one is measuring outcomes after the initiation of the program or not, whether a person is in the group due to participate in the program or not, and the interaction of these two indicator variables. Our *p*-values are based on heteroskedasticity-robust standard errors that also account for the clustering due to dependence in individuals' outcomes over time (so-called cluster-robust standard errors; see Cameron & Miller, 2015). Using a regression allows one to control for covariates.

Appendix Table D.3 shows that subjects who completed different waves of data collection do, in fact, differ in some ways. Most notable are differences in religiosity scores, and whether the subject came from an immigrant household. These differences arise because the waves that subjects completed varied mostly by program and peer-group locations, and the religiosity and immigrant household variables tend to be quite homogenous by location. We check for robustness to potential biases from such imbalances in two ways: (1) limiting the analysis to female subjects who complete *both* waves, and (2) controlling for covariates, including gender, age, the family religiosity

index, a dummy variable for whether or not the subject was a (first- or second-generation) immigrant to the community where they currently reside, and dummy variables for the communities where subjects currently reside. Table D.1 in Appendix D shows baseline characteristics for the sample used in the difference-in-differences analysis.

After estimating program effects, we also study structural relations between the ingroup censuring variable as an outcome and outgroup regard and psychological resources as regressors. To do so, we fit OLS models of the ingroup censuring index on the outgroup regard index and then the psychological resources index. We also examine how structural relations differ by ethnic group.

Finally, we bring the program effects estimates together with the structural analyses by estimating mediation effects. When considering the effect of the program on ingroup censuring, the mediation effect of outgroup regard is defined as the component of the total effect of the program that transmits via outgroup regard (Baron & Kenny, 1986; Imai et al., 2010). An analogous definition applies for the mediation effect of psychological resources. We estimate mediation effects for outgroup regard and psychological resources using a linear OLS mediation model (Imai, Keele, &, Yamamoto, 2010, pp. 57-58). Our mediation analyses use the basic specification from columns 1 in Table 2, given that the various control strategies do not appreciably change our estimates. We also estimate mediation effects separately by ethnicity. We follow current best practice and construct confidence intervals for the mediation effects using analytical methods that rely on a robust standard error estimator and the bootstrap, as recommended by Preacher and Hayes (2008).

Results

Effects of Program Participation

Figure 1 illustrates how the difference-in-differences estimate is constructed for the effect of the peace program on ingroup censuring. The wave 1 baseline mean value of the ingroup censuring index for participants is given by point A1, and the wave 1 baseline mean for non-participant peers is given by point B1. The wave 2 endline mean for participants is given by A2, and for non-participant peers it is given by B2. Recall that the key assumption is that the trends among the non-participant peers (the dashed line) describes how mean outcomes would have changed for participants in the absence of the program. Given this assumption, C shows what the participants' mean would have been at endline without the program. Then, the difference-in-differences estimate is given by A2 – C. The difference is about 0.4, indicating a substantial positive effect. This difference is measured in terms of ingroup-censuring index standard deviation units.

Table 2 displays results of regression analyses for the difference-in-differences effect that incorporate different types of controls and sample restrictions to check for robustness.[3] The first column shows the effect estimate that corresponds to what was presented in Figure 1---an effect size of 0.44 standard deviations. The estimated standard error is 0.24, yielding a marginally significant *p*-value of .07. (Note that this *p*-value is based on a two-sided test; had we applied a one-sided test, consistent with our directional hypothesis, the results would imply a *p*-value of .035.) The next three columns check the robustness of this estimate to various statistical controls. We first control for the

[3] Table D.6 in Appendix D shows that for Jewish subjects, the covariate controls do not exhibit strong relationships to any of the outcome variables, while for Arab-Palestinian subjects, gender is positively associated with all three outcomes, while age is positively associated with outgroup regard.

ethnicity of subjects (second results column) and then include the full set of covariate controls (third results column). Finally, the fourth results column shows estimates from a sample limited to female subjects who were present for both waves of data collection. In all cases, the estimated program effect remains very stable at about 0.4 standard deviations, although the standard error increases as we add more controls or reduce the sample size, and so the corresponding *p*-values rise as well.

[Table 2]

The last three results columns in Table 1 present results from an analysis of how ethnicity moderates these effects. For these three columns, the estimates in the "Peace program effect" row correspond to effects for Jewish subjects only. The row labeled "Arab-Palestinian moderation effect" shows coefficient estimates for a term that interacts the program effect with an indicator for whether the subject is Arab-Palestinian or not---that is, it yields the moderating interaction effect. A substantial interaction effect would imply substantial differences in the way the program affects Arab-Palestinian youth as compared to Jewish youth. The last two columns estimate the same effect but incorporate the covariate controls (second-to-last results column) and then restrict the sample to females present at both waves of data collection (last column). The estimated interaction effect is positive, but it is highly unstable across specifications. Moreover, the standard errors are very large. As such, we have no informative evidence of an interaction effect.

Table 3 shows results in the same format as Table 4 for the effects of the program on outgroup regard. To our surprise we do not see an effect of the program on outgroup regard in the pooled analysis ($b = 0.07$, $SE = 0.20$, $p = .72$). The estimate of a zero pooled effect holds when we control for ethnicity, include other covariate controls, and

then restrict ourselves to females that were present for both waves of data collection (results columns 2, 3, and 4). However, the pooled analysis masks important ethnic heterogeneity in the nature of the program's effects. These results are shown in the last three results columns of Table 3. For these columns, the estimates in the row labeled "Peace program effect" correspond to the effect for Jewish subjects. The estimates in the row labeled "Arab-Palestinian moderation effect" correspond to the moderator interaction effect, in which case the effect for Arab-Palestinian subjects is the sum of the coefficients in the "Peace program effect" and "Arab-Palestinian moderation effect" rows. Consistent with past research, we find evidence that the intervention increased Jewish subjects' outgroup regard. In the baseline specification (results column 5), we estimate an effect of .56 standard deviations (SE = 0.24, p = .02). The effect is even larger when we control for covariates (second-to-last column, b = 0.72, SE = 0.26, p = .01), although the estimate is smaller and far from statistically significant when we restrict the analysis to females present at both waves of data collection. However, the moderator interaction effect is always large, negative, and highly statistically significant. This implies that the program's effects were clearly different for Jewish as compared to Arab-Palestinian subjects. Indeed, the baseline estimate implies that the effect for Arab-Palestinian subjects is 0.56 – 1.21 = -0.65 standard deviations (SE = 0.28, p = .02). When we control for covariates and restrict the sample to females attending both waves of data collection, the resulting estimates are -0.87 (SE = 0.29, p = .004) and -0.72 (SE = 0.32, p = .03), respectively. Thus, the peace program *decreased* the outgroup regard of Arab-Palestinian subjects.

[Table 3]

Finally, Table 4 shows results in the same format as the previous two tables for the effects of the program on psychological resources. Here, we find no evidence of an effect. The coefficient estimates based on the data that pools both ethnic groups (results columns 1 through 4) are very close to zero, with correspondingly large *p*-values (in the .60 to .80 range). The estimates that incorporate the ethnicity interaction terms also yield results that are very close to zero, both for the coefficients that correspond to the effect on Jewish subjects (that is, the estimates in the row labeled "Peace program effects") and for the coefficients measuring the moderator interaction effect (in the row labeled "Arab-Palestinian moderation effect").

[Table 4]

To summarize, the peace program positively affects ingroup censuring, although this effect is only marginally significant. It positively affects outgroup regard for Jewish participants and negatively affects outgroup regard for Arab-Palestinian subjects. We find no effect on psychological resources.

Psychological Conditions of Ingroup censuring

We now turn to potential psychological processes that generate inclinations to censure ingroup members for aggressions toward the outgroup. For Jewish participants at least, the positive program effects for both ingroup censuring and outgroup regard suggest that high outgroup regard is a psychological condition for ingroup censuring. For Arab-Palestinian subjects, the results are not so clear. On the one hand, we have indication of a positive effect on ingroup censuring. However, the negative effect on outgroup regard goes against our expectations about how outgroup regard may mediate the effect on ingroup censuring. For both groups the null effect on psychological

resources suggests that this variable is not mediating the effect of the program on ingroup censuring, although we are still interested to know if there is any structural relation between the two variables. We thus examine structural relationships between ingroup censuring, outgroup regard, and psychological resources using all 149 subjects plus 65 repeat observations of those who participated twice (because of the difference-in-differences design). Given the group-based differences that our main analyses revealed, we conduct this analysis separately for each ethnic group. We also check if relationships are robust to controlling for the covariates used above (gender, age, religiosity, immigrant/resettlement status, and location) as well controlling for status as a peace program veteran.[4] Finally, we conduct mediation analyses relating these variables to the contact and sports peace program.

Table 5 shows results of OLS regressions of ingroup censuring on outgroup regard (panel A) and on psychological resources (panel B). We fit the models separately for Jewish and then Arab-Palestinian subjects. The last two columns fit the same model on the data pooling both ethnic groups with an interaction term meant to determine if differences between the two groups are statistically significant (as indicated by the results in the row labeled "Arab-Palestinian moderating effect").

The results reveal important asymmetries in the structural relationships. For Jewish youth, a standard deviation increase in the outgroup regard index corresponds to a 0.58 standard deviation increase in ingroup censuring ($p < .01$), but no such relationship

[4] As shown in Appendix D, Table D.5 veterans have higher average ingroup censuring, outgroup regard, and psychological resources index values, but these correlations are not robust to controlling for the other covariates.

exists for the Arab-Palestinian youth (for the difference between the coefficients, i.e. the Arab-Palestinian moderation effect presented in column 5, we have $p < .01$).

The reverse set of structural relationships holds for psychological resources. Table 5 shows a substantial relationship between changes in psychological resources and ingroup censuring for Arab-Palestinian subjects, and a weak relationship or, after controlling for background characteristics, essentially zero relationship for Jewish subjects. For Arab-Palestinian subjects, a standard deviation increase in the psychological resources factor score is associated with a 0.38 to 0.47 standard deviation increase in the ingroup censuring index ($p < .01$). The interaction of the Arab-Palestinian research subject indicator and psychological resources variable is not statistically significant at conventional levels ($p = .13$), meaning that the differences between the Arab-Palestinian and Jewish subjects should only be taken as suggestive in this case.

Our mediation analysis results are presented in Table 6. When we pool ethnic groups, we find no indication of mediation effects. However, for Jewish subjects, we find compelling evidence of a mediation effect for outgroup regard. Indeed the estimated mediation effect is as large (in face slightly larger) as the total effect. This suggests that the entirety of the effect of the program on ingroup censuring for Jewish subjects could be due to effects transmitted via outgroup regard. We find no such indication for Arab-Palestinian subjects. Nor is there any indication that psychological resources mediate the effect of the program, as we could have anticipated from the null effects of the program shown in Table 4.

This analysis of structural relations suggests that ingroup censuring, as a potential mechanism for mitigating inter-group conflict, has more to do with outgroup regard for

Jewish youth than for Arab-Palestinian youth. The opposite may be true for psychological resources, although such a structural relationship does not explain why inclinations to censure ingroup members increased among Arab-Palestinian participants.

Discussion

As predicted, the peace program increased our primary outcome of interest, ingroup censuring, although this effect is only marginally significant in a two-sided test. Our study is the first to show that intergroup contact interventions can affect the way participants evaluate ingroup peers' aggressions toward the outgroup, with implications for such participants' willingness to defend the interests of outgroup members. We found no overall program effect for outgroup regard and psychological resources. However, strong group-based differences emerged for outgroup regard. While the peace program positively affected outgroup regard among Jewish subjects, it negatively affected outgroup regard among Arab-Palestinian subjects.

The positive program effect for ingroup censuring among all participants and the positive program effect for outgroup regard among Jewish participants confirms that intergroup contact can sometimes be an effective tool for improving outgroup attitudes, even among members of groups in conflict. It is consistent with past research that found improvements in the attitudes of participants in co-ethnic sports programs in Israel. For example, following joint, year-long soccer activities Jewish youths participants had more positive attitudes by 11% - 28% across items in one study (N=322) and by 7% in a second study (N=151), and Arab-Palestinian youth's by 18.5% - 35% in the first study, and by 21-41% in the second study (Galily et al., 2013a, 2013b; Leitner et al., 2012;).

The negative program effect for outgroup regard among Arab-Palestinian participants is inconsistent with such past research. We note that the Arab-Palestinian participants in our study started with relatively high baseline outgroup attitudes. It is possible that the underprivileged Arab-Palestinian youth that made up the sample for our study were unusually positive towards Jewish Israelis at baseline as part of broader expectations about the program offering valuable resources such as extracurricular activities, sports equipment, and sports facilities. The reality of interacting with youth who belong to the outgroup may not have measured up to such high expectations. Disappointment may have led to a decline in outgroup regard. Of course, it is possible that this dip in outgroup regard after one year could be supplanted by positive attitude change in subsequent years. This is a question for future research.

Another possible explanation for the negative program effect is that Arab-Palestinian youth reacted negatively to the peace curriculum due to political incitement by Arab-Palestinian leaders. Such incitement is a form of authority disapproval, which can reduce or even reverse the contact-prejudice reduction link (Amir et al., 1980). If political incitement played a role we might see different program effects in Arab-Palestinian communities based on how active they are politically. Indeed, examining program effects by community revealed that the overall effect for Arab-Palestinian youth seems to be driven by one community (see Appendix D.4; community "G" in table A.1). Incidentally, this community is one of few Arab-Palestinian villages in Northern Israel with riots following the murder of Arab-Palestinian Muhammed Abu Khdeir in 2014 (Times of Israel Staff, 2014). While these riots suggest that youth may have been exposed to incitement in the prior year, we do not know what other, smaller-scale tensions may

have occurred in the other research communities and to what degree youth paid attention to any incitement. Future research should assess participant youth's awareness of such events and whether they perceive relevant authorities to support or object to the program. A final consideration is that, while the NGO we work with employs well-trained coaches, any intergroup contact intervention has the risk of creating negative contact experiences in addition to, or instead of, positive experiences. Negative contact has detrimental effects on outgroup attitudes (Barlow et al., 2013). Perhaps the experience of the Arab-Palestinian participants with the Jewish Israeli participants was more negative than vice versa in our study. Future research should assess the quality of participants' contact experiences in intergroup encounters to test this possibility.

Structural Relationships Between Variables

A further objective of our research was to explore the psychological conditions for ingroup censuring. For Jewish but not Arab-Palestinian participants we found a strong positive relationship between outgroup regard and ingroup censuring. Mediation analysis revealed that the positive program effect on ingroup censuring can largely be accounted for by increased outgroup regard among Jewish participants. This finding is consistent with the long-standing perspective that prejudice reduction is an important precondition for more peaceful and just societies (Allport, 1979). The mediation analysis for Jewish participants also suggests that while there is value in shifting ones' attention to intragroup outcomes, such as ingroup censuring, outgroup attitudes continue to matter as an important precondition for these outcomes.

For Arab-Palestinian, but not Jewish, participants we found a positive relationship between psychological resources and ingroup censuring. A reason for this asymmetry

may be in Arab-Palestinians' restricted access to material resources, which may make them more dependent on their own group (Austin-Smith & Fryer, 2005). Increased dependence on the ingroup would make defecting especially costly (Berman, 2011), and thus make participants especially sensitive to the availability of psychological resources. We did not find evidence that psychological resources mediated the positive program effect on ingroup censuring for Arab-Palestinians, given that the program did not affect psychological resources. However, several of the goals specified by the NGO that we worked with could be classified as psychological resource building goals, including the development of athletic skills and leadership capabilities. Building psychological resources may take a long time and a program effect on resources may emerge if we were to conduct long-term follow up data collection.

Limitations

The main limitation of our research is that our measure of ingroup censuring was private to prevent security risks for our subjects. Our theoretical presumption is that the effects we found would carry through into public behavior. Future research should investigate that possibility, perhaps by using observational data on public ingroup censuring.

A second limitation is that our sample consists of mostly adolescent girls. Maoz (2004) proposes that "confrontational" programs that incorporate discussions of the political context are more suitable for adults, while programs that focus on playful cooperation are more suitable for children and youth. Based on Maoz's (2004) analysis the peace program should be less effective for adults than youths. It is worth noting, however, that while discussing the conflict is not part of the official peace curriculum,

interviews with veteran participants revealed that they have intense and often emotional conversations about the conflict on the program sidelines. Tropp and Pettigrew (2005) found no evidence in their meta-analysis that gender moderates the contact-prejudice reduction link. For ingroup censuring, especially when it is public, we might expect stronger results for boys, since boys are often socialized to engage in heroic prosocial behaviors (Eagly & Crowley, 1986).

A limitation of our analysis of the peace program is that participants were not randomly assigned to participate. We matched the program and control group subjects as well as possible; nevertheless several differences emerged at baseline. For example, Arab-Palestinian control group subjects started off with less outgroup regard than Arab-Palestinian participants (Appendix Table D.1), but then these levels actually increased over time in the control group. This may be indicative of a "regression to the mean" effect and thus may partly account for the negative program effect we observed for Arab-Palestinians. A randomized controlled trial would be ideal to estimate unbiased program effect estimates.

In conclusion, our study is the first to show that intergroup contact interventions can increase participants' inclination to censure ingroup peers' aggressions toward the outgroup. For Jewish-Israeli participants this effect is mediated by improved attitudes. For Arab-Palestinian participants, this effect occurs despite a negative program effect on outgroup attitudes. Instead of attitudes, ingroup censuring is associated with resources in Arab-Palestinian participants. These findings, suggest that to promote peace, one needs to pay attention to processes that are specific to group members on the basis of their groups' status in society. One size does not fit all. To promote positive, conflict mitigating

behaviors like in-group censuring in all participants, such interventions should develop psychological resources in addition to promoting positive outgroup attitudes

References

Al Jazeera (2008, October 13). Acre faces fourth night of riots. Retrieved from http://www.aljazeera.com/news/middleeast/2008/10/200810112159349579 7.html

Allport, G. (1979). *The nature of prejudice.* Cambridge, MA: Perseus Books.

Angrist, J. D., & Pischke, J. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Amir, Y., Bizman, A., Ben-Ari, R., & Rivner, M. (1980). Contact between Israelis and Arabs: A theoretical evaluation of effects. *Journal of Cross-Cultural Psychology, 11*(4), 426-443.

Austen-Smith, D., & Fryer, R. G. Jr. (2005). An economic analysis of 'Acting White'. *The Quarterly Journal of Economics*, *120*(2), 551-583.

Avenanti, A., Sirigu, A., & Aglioti, S. M. (2010). Racial bias reduces empathic sensorimotor resonance with other-race pain. *Current Biology, 20*(11), 1018-1022.

Barlow, F. K., Paolini, S., Pedersen, A, Hornsey, M. J., Radke, H. R. M., Harwood, J., . . . Sibley, C. G. (2012). The contact caveat: Negative contact predicts increased prejudice more than positive contact predicts reduced prejudice. *Personality and Social Psychology Bulletin, 31*(38), 1629 – 1643.

Baron, R. M., & D. A. Kenny. (1986). The Moderator-Mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122–147.

Bassok, M., & Heruti-Sover, T. (2015, October 20). As wage gap between men and women widened, gap between Arabs and Jews shrunk. *Haaretz*. Retrieved from http://www.haaretz.com/israel-news/business/.premium-1.681336?=&ts=_1470836608744

Berman, E. (2011). *Radical, religious, and violent: The new economics of terrorism.* Cambridge, MA: MIT Press.

Bogardus, E. S. (1933). A social distance scale. *Sociology and Social Research, 17,* 265–271.

Brubaker, R., & Laitin, D. D. (1998). Ethnic and nationalist violence. *Annual Review of Sociology, 24*, 423–452.

Cameron, A. C., & D. L. Miller. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources 50*(2), 317-372.

Central Bureau of Statistics State of Israel (n.d.). Population and demography. Retrieved from http://www.cbs.gov.il/reader/?MIval=cw_usr_view_SHTML&ID=807

Central Bureau of Statistics State of Israel (2013). Israel in figures. Retrieved from http://www.cbs.gov.il/www/publications/isr_in_n13e.pdf

Central Bureau of Statistics State of Israel (2014). Statistical Abstracts of Israel, No. 65, Table 8.50. Retrieved from

http://www.cbs.gov.il/reader/shnaton/templ_shnaton_e.html?num_tab=st08_50&CYear=2014

Cikara, M., & Paluck, E. L. (2013). When going along gets you nowhere and the upside of conflict behaviors. *Social and Personality Psychology Compass*, 7(8), 559–571.

Crocker, J., Luhtanen, R., Blaine, B., & Broadnax, S. (1994). Collective self-esteem and psychological well-being among White, Black, and Asian college students. *Personality and Social Psychology Bulletin*, *20*(5), 503–513.

Dixon, J., Levine, M., Reicher, S., & Durrheim, K. (2012). Beyond prejudice: Are negative evaluations the problem and is getting us to like one another more the solution? *Behavioral and Brain Sciences*, *35*(06), 411–425.

Eagly, A., & Crowley, M. (1986). Gender and helping behavior: A meta-analytical review of the social psychological literature. *Psychological Bulletin, 100*(3), 283-308.

Eliezer, D., & Major, B. (2012). It's not your fault: The social cost of claiming discrimination on behalf of someone else. *Group Processes and Intergroup Relations, 15*(4), 487-502.

Fearon, J. D., & Laitin, D. D. (1996). Explaining interethnic cooperation. *The American Political Science Review, 90*(4), 715–735.

Football for Hope (2014). Retrieved from http://www.fifa.com/aboutfifa/socialresponsibility/footballforhope/

Galily, Y., Leitner, M., & Shimon, P. (2013a). Coexistence and sport: The Israel case. *Journal of Comparative Research in Anthropology and Sociology*, *4*(2), 181-194.

Galily, Y., Leitner, M., & Shimon, P. (2013b). The effects of three Israeli sports programs on attitudes of Arabs and Jews toward one another. *Journal of Aggression, Conflict, and Peace Research, 5*(4), 243-358.

Heatherton, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology*, *60*(6), 895–910.

Hammack, P. L. (2006). Identity, conflict, and coexistence: Life stories of Israeli and Palestinian Adolescents. *Journal of Adolescent Research, 21*(4), 323-369.

Hameiri, B., Bar-Tal, D., & Halperin, E. (2014). Challenges for peacemakers: How to overcome socio-psychological barriers. *Policy Insights from the Behavioral and Brain Sciences, 1*(1), 164-171.

Heinrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2-3), 1-75.

Huesmann, R. L., & Guerra, N. G. (1997). Children's normative beliefs about aggression and aggressive behavior. *Journal of Personality and Social Psychology*, 72(2), 408–419.

Imai, K. L. Keele, & T. Yamamoto. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science, 25*(1), 51-71.

Kydd, A., & Walter, B. F. (2002). Sabotaging the peace: The politics of extremist violence. *International Organization*, *56*(2), 263–296.

Lazarus, R. S., & DeLongis. A. (1983). Psychological stress and coping in aging. *American Psychologist*, *38*(3), 245–254.

Leitner, M. J., Galily, Y., & Shimon, P. (2012). The effects of Peres Center for Peace Sports Programs on the attitudes of Arab and Jewish Israeli Youth. *Leadership and Policy Quarterly*, *1*(2), 109-121.

Livtak-Hirsch, T., Galily, Y., & Leitner, M. (2016). Evaluating conflict mitigation and health improvement through soccer: a two-year study of 'Mifalot's United Soccer for Peace' programme. *Soccer & Society*, *17*(2), 209-224.

Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin*, *18*(3), 302–318.

Major, B., & O'Brien, L. T. (2005). The social psychology of stigma. *Annual Review of Psychology*, *56*(1), 393–421.

Maoz, I. (2000a). Power relations in inter-group encounters: A case study of Jewish-Arab encounters in Israel. *International Journal of Intercultural Relations*, *24*(4), 259-277.

Maoz, I. (2000b). Multiple conflicts and competing agendas: A framework for conceptualizing structured encounters between groups in conflict – The case of a coexistence project of Jews and Palestinians in Israel. *Peace and Conflict: Journal of Peace Psychology*, *6*(2), 135-156.

Maoz, I. (2004). Coexistence is in the eye of the beholder: Evaluating intergroup encounter interventions between Jews and Arabs in Israel. *Journal of Social Issues, 60* (2), 437-452.

Maoz, I. (2012). The dangers of prejudice reduction interventions: Empirical evidence from encounters between Jews and Arabs in Israel. *Behavioral and Brain Sciences, 35*(6), 441-442.

Möller, I., & Krahé, B. (2009). Exposure to violent video games and aggression in German adolescents: A longitudinal analysis. *Aggressive Behavior, 35*(1), 75–89.

Noor, M., Shnabel, N., Halabi, S., & Nadler, A. (2012). When suffering begets suffering the psychology of competitive victimhood between adversarial groups in violent conflicts. *Personality and Social Psychology Review, 16*(4), 351-374.

Paolini, S., Hewstone, M., Cairns, E., & Voci, A. (2004). Effects of direct and indirect cross-group friendships on judgments of Catholics and Protestants in Northern Ireland: The mediating role of an anxiety-reduction mechanism. *Personality and Social Psychology Bulletin, 30*(6), 770–786.

Parrillo, V. N., & Donoghue C. (2005). Updating the Bogardus social distance studies: A new national survey. *The Social Science Journal, 42*(2), 257–271.

Peacebuilding and Reconcliation (2014). Retrieved from: http://www.usaid.gov/what-we-do/working-crises-and-conflict/peacebuilding-and-reconciliation

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology, 49*(1), 65-85.

Pettigrew, T. F., & Tropp, L. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology, 90*(5), 751-783.

Preacher, K.J., & A.F. Hayes. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods 40*(3), 879-891.

Rousseau, D. (2014, August 18). Jewish-Muslim wedding brings out bitter Israel tensions. *Agence France Press*, online edition. Retrieved from http://news.yahoo.com/jewish-muslim-wedding-brings-bitter-israel-tensions-151731066.html

Saguy, T., Chernyak-Hai, L., Andrighetto, L., & Bryson, J. (2013). When the powerful feel wronged: The legitimization effects of advantaged group members' sense of being accused for harboring racial or ethnic biases. *European Journal of Social Psychology*, *43*(4), 292-298.

Saguy, T., Tausch, N., Dovidio, J. F., & Pratto, F. (2009). The irony of harmony: Intergroup contact can produce false expectations for equality. *Psychological Science*, 20(1), 114-121.

Schroeder, J., & Risen, J. L. (2016). Befriending the enemy: Outgroup friendship longitudinally predicts intergroup attitudes in a coexistence program for Israelis and Palestinians. *Group Processes and Intergroup Relations*, *19*(1), 72-93.

Sherer, M., & Adams, C. H. (1983). Construct validation of the self efficacy scale. *Psychological Reports*, *53*(3), 899–902. doi:10.2466/pr0.1983.53.3.899

Times of Israel Staff (2014, July 6). Protests erupt again in East Jerusalem, Southern Israel. *Times of Israel*. Retrieved from http://www.timesofisrael.com/spate-of-riots-erupt-in-northern-israel/

Tropp, L., & Pettigrew, T. F. (2005). Relationships between intergroup contact and prejudice among minority and majority status groups. *Psychological Science, 16*(12), 951-957.

Zonszein, M. (2014, September 26). How Israel silences dissent. *The New York Times.* Retrieved from *http://www.nytimes.com/2014/09/27/opinion/how-israel-silences-dissent.html*.

Tables

Table 1

Research Subject Characteristics

Jewish (N=114 observations, including 77 subjects plus repeated pre-post observations for 37 subjects)

Categorical variables	Yes (N)	(%)	No (N)	(%)
Female	96	84%	18	16%
Immigrant household	38	33%	76	67%
Peace program participant	59	52%	55	48%
Veteran peace program participant	16	14%	98	86%
Continuous variables	Mean	S.D.	Min.	Max.
Age	13.46	1.97	10	21
Religiosity index (1=most religious, 4=least)	3.20	0.77	2	4
Outgroup feeling thermometer scale	45.24	24.75	0	100
Social distance scale	3.32	1.15	1	5
Outgroup regard index	-0.12	1.07	-2.34	1.58
Collective self esteem scale	5.57	0.90	1	7
Self-efficacy scale	4.13	0.60	1.71	5
State self-esteem scale	5.21	1.19	1.45	6.85
Psychological resources index	-0.14	1.11	-3.86	1.33
Ingroup censuring index	3.67	0.61	1	4

Arab-Palestinian (N=100 observations, including 72 subjects plus repeated pre-post observations for 28 subjects)

Categorical variables	Yes (N)	(%)	No (N)	(%)
Female	86	86%	14	14%
Immigrant household	24	24%	76	76%
Peace program participant	58	58%	42	42%
Veteran peace program participant	23	23%	77	77%
Continuous variables	Mean	S.D.	Min.	Max.
Age	12.89	2.16	10	22
Religiosity index (1=most religious, 4=least)	2.29	0.80	1	4
Outgroup feeling thermometer scale	48.60	30.07	0	100
Social distance scale	3.68	1.00	1	5
Outgroup regard index	0.14	0.89	-2.34	1.55
Collective self esteem scale	5.60	1.08	1	7
Self-efficacy scale	3.95	0.50	2.65	4.88
State self-esteem scale	5.75	0.88	2.85	7
Psychological resources index	0.15	0.84	-2.23	1.36
Ingroup censuring index	3.65	0.54	1	4

Note. The table shows how the data for Jewish and Arab-Palestinian subjects are distributed over key background characteristics. Statistics are calculated using the total set of observations, which includes a combination of single-wave and repeated observations, as indicated in the header to each panel. Incoming participants, their non-participant peers, and veteran participants are included. For a table of incoming program participant and non-participant peer characteristics at baseline, see Appendix Table A.4.

Table 2

Difference-in-Differences Estimates of Program Effects on Ingroup censuring

	Pooled	Pooled, with ethnic controls	Pooled, with all covariates	Both waves, female only	Ethnic interaction effect	Ethnic interaction effect, with covs.	Ethnic interaction effect, both waves
	b/se/p	b/se/p	b/se/p	b/se/p	b/se/p	b/se/p	b/se/p
Peace program effect (diff.-in.diff.)	0.44	0.40	0.42	0.39	0.29	0.38	0.12
	(0.24)	(0.23)	(0.26)	(0.27)	(0.21)	(0.21)	(0.22)
	[0.07]	[0.09]	[0.10]	[0.16]	[0.17]	[0.07]	[0.59]
Arab-Palestinian average diff.		0.31	0.67		0.35	0.66	0.83
		(0.29)	(1.05)		(0.29)	(1.06)	(0.40)
		[0.29]	[0.53]		[0.23]	[0.54]	[0.04]
Arab-Palestinian moderation effect (interaction term)					0.24	0.10	0.71
					(0.51)	(0.59)	(0.55)
					[0.63]	[0.87]	[0.20]
R-squared	0.03	0.08	0.13	0.05	0.08	0.13	0.12
Number of observations (including repeated observations)	175	175	175	104	175	175	104

Note.
OLS regression coefficients.
Cluster-robust standard errors in parentheses accounting for repeated subjects.
Two-sided p-values in brackets.
Dependent variable is ingroup censuring index, standardized relative to the pooled mean and standard deviation.
Coefficients are in outcome standard deviation units.
Specifications with covariates control for gender, immigrant status, age, religiosity, and location dummy variables.
By "pooled" we mean that the estimates apply to both Jewish and Arab-Palestinian subjects.
For the last three columns, the "Peace program effect" estimates correspond to Jewish subjects only. Estimates for Arab-Palestinian subjects can be computed from the sum of the "Peace program effect" coefficient and the "Arab-Palestinian moderation effect" coefficient.

Table 3

Difference-in-Differences Estimates of Program Effects on Outgroup regard

	Pooled	Pooled, with ethnic controls	Pooled, with all covariates	Both waves, female only	Ethnic interaction effect	Ethnic interaction effect, with covs.	Ethnic interaction effect, both waves
	b/se/p	b/se/p	b/se/p	b/se/p	b/se/p	b/se/p	b/se/p
Peace program effect (diff.-in.diff.)	0.07	0.04	0.06	-0.10	0.56	0.72	0.29
	(0.20)	(0.19)	(0.21)	(0.22)	(0.24)	(0.26)	(0.25)
	[0.72]	[0.83]	[0.78]	[0.63]	[0.02]	[0.01]	[0.25]
Arab-Palestinian average diff.		0.01	0.61		-0.20	0.71	0.16
		(0.29)	(0.64)		(0.30)	(0.63)	(0.45)
		[0.96]	[0.35]		[0.51]	[0.27]	[0.72]
Arab-Palestinian moderation effect (interaction term)					-1.21	-1.59	-1.01
					(0.37)	(0.39)	(0.40)
					[0.00]	[0.00]	[0.02]
R-squared	0.03	0.05	0.23	0.03	0.07	0.26	0.10
Number of observations (including repeated observations)	175	175	175	104	175	175	104

Note.
OLS regression coefficients.
Cluster-robust standard errors in parentheses accounting for repeated subjects.
Two-sided p-values in brackets.
Dependent variable is outgroup regard index.
Coefficients are in outcome standard deviation units.
Specifications with covariates control for gender, immigrant status, age, religiosity, and location dummy variables.
By "pooled" we mean that the estimates apply to both Jewish and Arab-Palestinian subjects.
For the last three columns, the "Peace program effect" estimates correspond to Jewish subjects only. Estimates for Arab-Palestinian subjects can be computed from the sum of the "Peace program effect" coefficient and the "Arab-Palestinian moderation effect" coefficient.

Table 4

Difference-in-Differences Estimates of Program Effects on Psychological Resources

	Pooled	Pooled, with ethnic controls	Pooled, with all covariates	Both waves, female only	Ethnic interaction effect	Ethnic interaction effect, with covs.	Ethnic interaction effect, both waves
	b/se/p	b/se/p	b/se/p	b/se/p	b/se/p	b/se/p	b/se/p
Peace program effect (diff.-in.diff.)	0.04	-0.06	0.07	-0.18	-0.01	0.08	-0.23
	(0.21)	(0.20)	(0.21)	(0.20)	(0.26)	(0.28)	(0.26)
	[0.87]	[0.78]	[0.75]	[0.38]	[0.98]	[0.78]	[0.38]
Arab-Palestinian average diff.		0.63	0.64		0.61	0.64	0.71
		(0.26)	(0.84)		(0.26)	(0.84)	(0.34)
		[0.02]	[0.45]		[0.02]	[0.45]	[0.04]
Arab-Palestinian moderation effect (interaction term)					-0.12	-0.03	0.03
					(0.42)	(0.42)	(0.43)
					[0.77]	[0.94]	[0.94]
R-squared	0.07	0.15	0.26	0.11	0.15	0.26	0.21
Number of observations (including repeated observations)	175	175	175	104	175	175	104

Note.
OLS regression coefficients.
Cluster-robust standard errors in parentheses accounting for repeated subjects.
Two-sided p-values in brackets.
Dependent variable is psychological resources index.
Coefficients are in outcome standard deviation units.
Specifications with covariates control for gender, immigrant status, age, religiosity, and location dummy variables.
By "pooled" we mean that the estimates apply to both Jewish and Arab-Palestinian subjects.
For the last three columns, the "Peace program effect" estimates correspond to Jewish subjects only. Estimates for Arab-Palestinian subjects can be computed from the sum of the "Peace program effect" coefficient and the "Arab-Palestinian moderation effect" coefficient.

Table 5

Structural Relationship Between Ingroup censuring (Outcome) and then Outgroup regard (Panel A) and Psychological Resources (Panel B)

Panel A: Outgroup regard and ingroup censuring

	Jewish-Israeli	Jewish-Israeli w/ covs.	Arab-Palestinian	Arab-Palestinian w/ covs.	Pooled	Pooled w/ covs.
	b/se/p	b/se/p	b/se/p	b/se/p	b/se/p	b/se/p
Outgroup regard index	0.58	0.58	-0.04	-0.17	0.58	0.57
	(0.15)	(0.16)	(0.13)	(0.15)	(0.15)	(0.15)
	[0.00]	[0.00]	[0.78]	[0.26]	[0.00]	[0.00]
Arab-Palestinian moderation effect (interaction term)					-0.62	-0.67
					(0.20)	(0.21)
					[0.00]	[0.00]
R-squared	0.35	0.40	0.01	0.18	0.21	0.29
N	114	114	100	100	214	214

Panel B: Psychological resources and ingroup censuring

	Jewish-Israeli	Jewish-Israeli w/ covs.	Arab-Palestinian	Arab-Palestinian w/ covs.	Pooled	Pooled w/ covs.
	b/se/p	b/se/p	b/se/p	b/se/p	b/se/p	b/se/p
Psych. resources index	0.19	0.06	0.47	0.38	0.19	0.04
	(0.12)	(0.16)	(0.14)	(0.13)	(0.12)	(0.15)
	[0.11]	[0.72]	[0.00]	[0.00]	[0.11]	[0.78]
Arab-Palestinian moderation effect (interaction term)					0.28	0.35
					(0.18)	(0.19)
					[0.13]	[0.07]
R-squared	0.04	0.15	0.18	0.26	0.10	0.18
N	114	114	100	100	214	214

Notes for both panels.

Dependent variable is ingroup censuring index (standardized) in both panels.

Cluster robust standard errors in parentheses accounting for repeated subjects.

Two-sided p-values in brackets.

The first row of results in each panel shows standardized coefficients, measuring a standard deviation change in outcome given standard deviation change in regressor.

Specifications with covariates control for gender, veteran, immigrant status, age, religiosity, and location dummy variables.

For the last two columns, the first row estimates correspond to Jewish subjects only. Estimates for Arab-Palestinian subjects can be computed from the sum of the first row coefficient and the "Arab-Palestinian moderation effect" coefficient.

Table 6

Mediation Effects of Outgroup regard and Psychological Resources on Ingroup censuring

	Pooled	Jewish	Arab-Palestinian
Total effect of the peace program	0.44	0.29	0.54
95% confidence interval (analytical)	(-0.15, 1.07)	(-0.34, 1.07)	(-0.40, 1.45)
95% confidence interval (bootstrap)	(-0.17, 1.09)	(-0.36, 1.09)	(-0.42, 1.46)
Outgroup regard mediation effect	0.03	0.34	0.03
95% confidence interval (analytical)	(-0.12, .20)	(0.06, 0.75)	(-0.18, 0.26)
95% confidence interval (bootstrap)	(-0.13, 0.21)	(0.06, 0.76)	(-0.17, 0.26)
Psychological resources mediation effect	0.01	0.00	-0.06
95% confidence interval (analytical)	(-0.10, 0.13)	(-0.09, 0.09)	(-0.39, 0.28)
95% confidence interval (bootstrap)	(-0.10, 0.13)	(-0.08, 0.09)	(-0.39, 0.28)
N	175	98	77

Note.
Outcome variable is the ingroup censuring index (standardized).
Linear, ordinary least squares mediation analysis results.
The total effect corresponds to the estimates displayed in Table 2.
The mediation effect is the component of the total effect that transmits through the mediator (i.e., total effect = mediation effect + residual effect).
The analytical confidence intervals are derived from cluster-robust standard error estimates.
The bootstrap confidence intervals are obtained from a cluster-bootstrap.
For both, we cluster by individuals to account for over-time dependence.

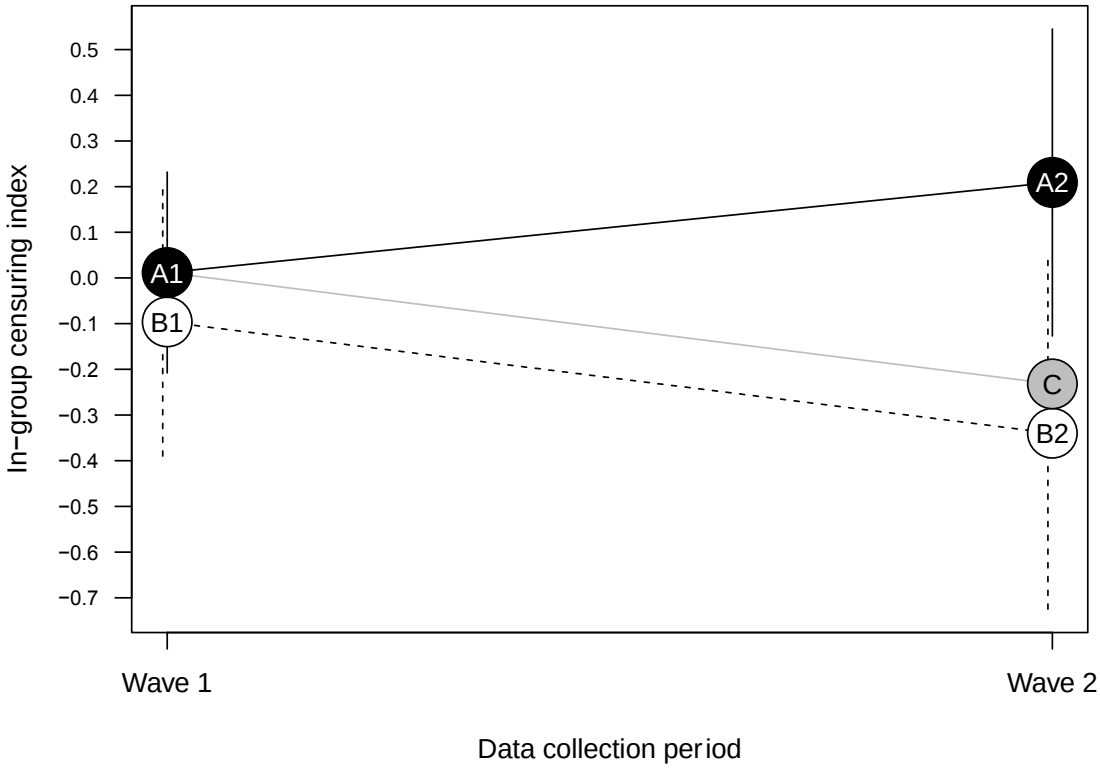

Figure 1. Changes in ingroup censuring index means over time for participants and non-participants, and corresponding difference-in-difference estimation. The graph displays how mean levels of the ingroup censuring index changed from wave 1 to wave 2 for participants (black points and black lines) and non-participant peers (white points and dashed lines). The horizontal bars show 95% confidence intervals. The wave 1 baseline mean for participants is given by A1, for non-participant peers it is B1. The wave 2 endline mean for participants is given by A2, for non-participant peers, it is B2. The gray point labeled as C is the difference-in-differences estimate of the counterfactual mean for participants---that is, what the mean level would have been had there been no program. The difference-in-differences program effect is estimated as A2-C.